基于改进YOLACT的油茶叶片炭疽病感染严重程度分级模型

聂刚刚[1,2], 饶洪辉[1,2*], 李泽锋[1,2], 刘木华[1,2]

(1. 江西农业大学 工学院,江西南昌 330045,中国; 2. 江西省现代农业装备重点实验室,江西 南昌 330045,中国)

摘　要：［**目的/意义**］炭疽病（*anthracnose*）作为油茶生长过程中重要的病害，其严重程度的精准判定对于精准施药和科学管理具有重大意义。本研究提出了一种改进YOLACT（You Only Look At CoefficienTs）分级模型Camellia-YOLACT，旨在实现对油茶叶片炭疽病感染严重程度的自动、高效判定。［**方法**］首先在YOLACT主干网络部分使用Swin-Transformer来进行特征提取。Transformer架构的自注意力机制拥有全局感受野及移位窗口等特性，有效地增强了模型的特征提取能力；引入加权双向特征金字塔网络，融合不同尺度的特征信息，加强模型对不同尺度目标的检测能力，提高模型的检测精度；在激活函数的选择上，采用非线性能力更强的HardSwish激活函数替换原模型的ReLu激活函数。由于HardSwish在负值区域不是完全截断，对于输入数据中的噪声具有更高的鲁棒性，自然环境下的图像有着复杂的背景和前景信息，HardSwish的鲁棒性有助于模型更好地处理这些情况，进一步提升精度。［**结果和讨论**］采用迁移学习方式在油茶炭疽病感染严重程度分级数据集上进行实验验证。消融实验结果表明，本研究提出的Camellia-YOLACT模型的mAP_{75}为86.8%，较改进前提升5.7%；mAP_{all}为78.3%，较改进前提升2.5%；mAR为91.6%，较改进前提升7.9%。对比实验结果表明，Camellia-YOLACT在精度和速度方面表现均好于SOLO（Segmenting Objects by Locations），与Mask R-CNN算法相比，其检测速度提升了2倍。在室外的36组分级实验中进一步验证了Camellia-YOLACT模型的性能，其对油茶炭疽病严重程度的分级正确率达到了94.4%，K值平均绝对误差为1.09%。［**结论**］本研究提出的Camellia-YOLACT模型在油茶叶片和炭疽病病斑分割上具有较高的精度，能够实现对油茶炭疽病严重程度的自动分级，为油茶病害的精准防治提供技术支持，进一步推动油茶炭疽病诊断的自动化和智能化。

关键词：油茶；叶部病害；炭疽病；BiFPN；YOLACT；Transformer；深度学习

中图分类号：TP391.4；S431.9　　**文献标志码：**A　　**文章编号：**SA202402002

引用格式：聂刚刚, 饶洪辉, 李泽锋, 刘木华. 基于改进YOLACT的油茶叶片炭疽病感染严重程度分级模型[J]. 智慧农业(中英文), 2024, 6(3): 138-147. DOI：10.12133/j.smartag.SA202402002

NIE Ganggang, RAO Honghui, LI Zefeng, LIU Muhua. Severity Grading Model for *Camellia Oleifera* Anthracnose Infection Based on Improved YOLACT[J]. Smart Agriculture, 2024, 6(3): 138-147. DOI：10.12133/j.smartag.SA202402002 (in Chinese with English abstract)

0　引　言

油茶是世界四大木本油料植物之一，主要分布在中国南方各省（区）[1]。油茶病害是导致茶油品质下降和农民经济损失的主要原因之一。其中油茶炭疽病（*anthracnose*）是油茶林非常普遍且危害最严重的病害，每年由于油茶炭疽病而造成油茶籽实减产10%～30%，重病区减产40%～50%[2]。为了帮助农民采取针对性的治疗措施和灵活的农药用量策略，高效、精准地量化油茶炭疽病的严重程度非常必要。传统的油茶炭疽病严重程度评估主要依靠有经验的生产者或专业植保人员在茶园人工逐个观察叶片病斑，费时费力且存在滞后性。近些年，计算机视觉和深度学习在诊断植物病害方面取得了广泛应用，为油茶炭疽病严重程度分级提供了新的解

收稿日期：2024-02-02

基金项目：江西省科技计划项目（20141BBF60057）；江西省林业厅油茶研究专项（YCYJZX2023221）

作者简介：聂刚刚，研究方向为机器视觉。E-mail：nieganggang1997@163.com

*通信作者：饶洪辉，博士，教授，研究方向为现代农业装备、机器视觉。E-mail： rhh58@sohu.com

决方案。

传统的图像分割算法如阈值算法、分水岭算法和GraphCut等算法依赖人工设计特征[3]，需要图像颜色、灰度、形状和纹理等低层次特征与形态学操作相结合分割出病斑，分割精度不高，泛化能力较差，对于复杂背景图像适应性较差。而基于深度学习方法的分割算法则能够自动从原始图像中学习和提取高级别的特征表示，有效地分割出图像中目标对象的不同区域，即便是在复杂的背景下也能保持较好的分割性能，尤其适用于对病害分布不均匀、形态多变的油茶叶部病害图像的分割与严重程度分级。

基于深度学习的植物病害严重程度评估的方法主要分为基于分类的分级方法和基于分割的分级方法。基于分类的分级方法通常将植物的病害程度划分为前期、中期及后期等几个区间，将病害程度分级问题转化为图像分类问题[4]。Prabhakar等[5]将公开数据集PlantVillage中的番茄叶部图像划分为健康、轻度、中度和重度，使用ResNet101模型进行训练，该模型对番茄叶部病害的分级准确率达到94.6%。Tendang和Chamnongthai[6]训练了一个深度卷积神经网络模型，该模型识别水稻叶部病害早期、中期和晚期的准确率分别为96.40%、96.40%和96.56%。万军杰等[7]采用迁移学习技术与GoogLeNet模型对6种果园作物的25类病虫害样本进行识别与危害等级分级研究，危害程度分级精度达到92.78%。LIU等[8]提出了一种改进的炭疽病、褐斑病、螨虫、黑腐病、霜霉病、叶枯病诊断模型DICNN，整体准确率为97.22%。与GoogLeNet和ResNet-34相比，识别准确率分别提高了2.97%和2.55%，为深度学习在病虫害诊断领域的应用奠定了理论基础。

基于分类的分级方法受制于分类网络模型的局限性，需要大量的样本图像供模型学习不同分级下的病害特征，不同类别间分界阈值模糊难以量化，而且对数据集的病害级别标注的准确率有很高的要求，也不能根据诊断标准的变化调整病害程度的诊断结果，实际应用价值有限。

相对于基于分类的方法，基于分割的分级方法不需要预先设定固定的病害级别区间，而是通过病斑面积占叶片总面积的比例来定量评价病害的严重程度，这种方法更为客观和连续，能够捕捉到病害细微程度的变化[9]。Garg等[10]使用语义分割网络，同时分割单个叶片实例和相应的染病区域，并将该模型部署在无人机上实现了对植物病害程度的自动量化。Goncalves等[11]训练了六种语义分割网络来对大豆锈病和小麦棕褐色斑病的病变叶片进行分割对比实验，结果表明特征金字塔网络（Feature Pyramid Networks, FPN）、U-Net和DeepLabv3+表现优异，将病斑分割结果用于衡量病害严重程度获得了较好效果。茹佳棋等[12]提出了一种基于改进UNet++的葡萄黑腐病病斑分割模型，使用自适应软阈值细化病斑边缘，通过多尺度特征融合聚合低层特征图高分辨率信息恢复小病斑特征，该模型分级准确率达97.41%。邓朝等[13]采用Mask R-CNN（Mask Region-based Convolutional Neural Network）算法对马铃薯叶片晚疫病进行量化评价，准确率为87.5%。以上研究表明相比于传统的图像处理方法和基于分类网络的方法，使用分割网络进行病害程度分级有着更好的效果，然而以上研究大多基于语义分割网络UNet和二阶段实例分割网络Mask R-CNN开展，由于网络架构原因，存在着模型参数量大，推理速度较慢等缺点，难以满足在茶园中的实时应用。因此，研究一种更高效的自然环境下油茶叶部病害严重程度分级的方法十分有必要。

本研究以自然环境下油茶炭疽病为研究对象，提出一种基于改进YOLACT（You Only Look At CoefficienTs）的油茶炭疽病分割模型Camellia-YOLACT，并实现炭疽病严重程度分级的方法。在单阶段实例分割网络YOLACT基础上，引入Swin-Transformer[14]模型作为主干特征提取网络，使用BiFPN[15]作为特征融合网络，更好地融合不同尺度的特征信息，同时对激活函数进行优化，以提高模型准确性与实时性，实现对叶片和病斑区域的精准分割，最后按照分割的病斑面积占分割的叶片面积比例进行病害严重程度分级，该方法为油茶炭疽病严重程度分级与灵活施药提供技术支持。

1 材料与方法

1.1 实验数据

1.1.1 图像数据采集

油茶炭疽病图像采集于江西省林业科学院国家油茶良种基地（28°74′N，115°82′E）。采集时间为2023年3月至4月，为贴近实际油茶园复杂环境下的应用，采集时保留背景噪声，包含不同光照条件和不同病害程度图像样本，采集设备为华为nova6se智能手机，图像分辨率像素为3 000×3 000。人工检查后剔除模糊图像、叶部遭到遮挡或截断的

图像以及相似度过高的图像，经筛选后最终得到有效数据图像共计975张。为提高模型训练效率，将原始图像像素统一调整为544×544，使用Labelme软件对图像中的叶片和病斑区域进行标注，结果如图1所示，最终构建油茶炭疽病严重程度分级数据集。



a.原始图像 b.标注后图像

图1 油茶炭疽病数据集Labelme标注样例

Fig. 1 Samples of *Camellia oleifera* leaf anthracnose dataset label by Labelme

1.1.2 图像数据增强

为了避免训练过程中出现过拟合现象，提高模型鲁棒性和泛化能力，采用图像数据增强方式对数据集进行扩充。数据增强策略分别为水平翻转、垂直翻转、随机亮度、随机对比度、高斯噪声，增强后样本如图2所示，最终样本数共计5 850张，按照7∶2∶1比例划分为训练集4 095张、验证集1 170张和测试集585张。

1.1.3 油茶炭疽病严重程度分级

植物叶部病害的严重程度通常用病斑区域面积与叶片总面积的百分比*K*来衡量。使用分割网络可以得到病斑区域和叶片区域各自的掩膜图，然后分别计算病斑区域和叶片区域的面积，二者之比即可评估油茶炭疽病的严重程度，病害严重程度计算如公式（1）所示。

$$K = \frac{S_{\text{Disease}}}{S_{\text{Leaf}}} \times 100\% \tag{1}$$

式中：S_{Disease}为病斑区域面积，pixel；S_{Leaf}为叶片总面积，pixel；*K*为病斑占叶片比例。

参照DB34/T 3863—2021《茶炭疽病测报调查与防治技术规程》[16]中茶炭疽病严重度分级标准具体分级，按照病斑占叶片面积的比例分为5级，如表1所示。



a. 原始图像 b. 水平翻转 c. 垂直翻转

d. 随机亮度 e. 随机对比度 f. 高斯噪声

图2 油茶炭疽病数据集扩充示例

Fig.2 Example of *Camellia oleifera* leaf anthracnose dataset expansion

表1 油茶炭疽病严重程度分级标准

Table 1 Classification standard for the degree of *Camellia oleifera* leaf anthracnose

严重程度分级	分级标准
0级	无病斑
1级	$K \leqslant 25\%$
2级	$25\% < K \leqslant 50\%$
3级	$50\% < K \leqslant 75\%$
4级	$K > 75\%$

注：*K*为油茶炭疽病严重程度分级系数。

1.2 YOLACT的分割模型改进

传统的实例分割模型如Mask R-CNN[17]为二阶段模型，通常先检测后分割，即先通过目标检测方法找出实例区域，再在边界框内进行语义分割，最后每个分割结果作为不同的实例输出，该方法检测和分割整体耗时过长，难以实现实时检测，实用性不足。YOLACT[18]是一个单阶段的用于实时实例分割的模型，同时进行检测与分割，在速度上表现优异。

YOLACT网络结构主要由特征提取主干网络（Feature Backbone）、特征金字塔网络（Feature Pyramid Net, FPN）、原型生成分支（Protonet），以及检测头分支（Prediction Head）组成。其中原型生成分支用来生成原型掩膜（Prototype Mask），检测头分支用来预测每个实例的掩膜系数（Mask Co-

efficients），两个分支并行执行，最后将原型掩膜和掩膜系数线性组合生成实例掩膜。YOLACT模型结构轻量化，推理速度快，模型参数量少，但存在精度较差的问题。为在保证检测速度的前提下提高对油茶炭疽病叶部区域的分割精度，对YOLACT模型进行改进。首先将特征提取主干网络替换为窗口自注意力网络Swin-Transformer，利用自注意力机制来捕获全局的上下文信息，提升模型的特征提取能力。然后使用特征融合效果更优的加权双向特征金字塔网络（Bidirectional Feature Pyramid Network, BiFPN）来融合不同尺度的特征信息，提高检测精度。最后将原型生成分支和检测头中的ReLu激活函数替换为HardSwish激活函数提升模型表达能力。改进后的模型为Camellia-YOLACT，结构如图3所示。

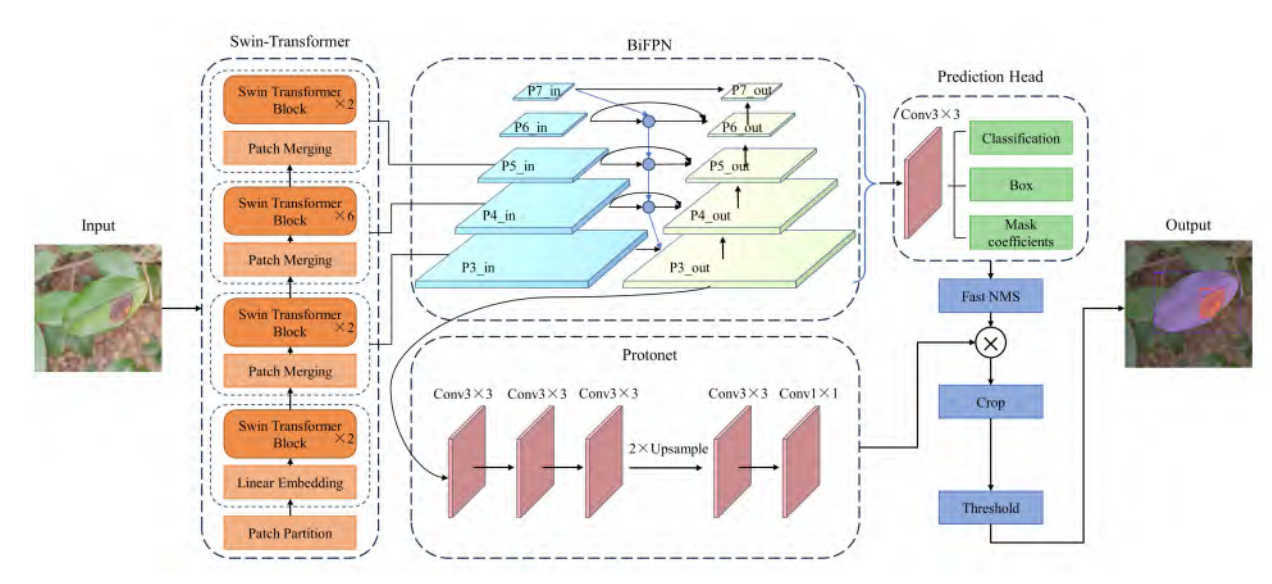


图3 Camellia-YOLACT模型结构

Fig.3 Structure of the Camellia-YOLACT model

1.2.1 特征提取主干网络Swin-Transformer

原始的YOLACT模型采用ResNet50[19]作为特征提取主干网络，由于卷积神经网络（Convolutional Neural Networks, CNN）仅关注小范围的局部特征，全局特征提取能力差。使用Swin-Transformer作为特征提取主干网络，可以捕获全局和局部上下文信息，扩大感受野，增强网络的特征提取能力。因此，引入与ResNet50复杂度相当的Swin-T版本作为YOLACT的主干网络。Swin-Transformer模型架构如图4所示，使用了层次化构建方式，构建了4个Stage，每个Stage都对特征图进行下采样操作进而得到多尺度的特征图。Swin-Transformer首先会使用4×4的窗口对输入图像进行图块分割（Patch Partition），之后沿通道方向将分割的图块进行展平，接着在Stage1中利用线性嵌入（Linear Embeding）来调整向量维度，随后进入Swin Transformer模块进行自注意力计算提取图像特征。图块合并（Patch Merging）将邻近图块拼合在一起完成下采样，此时图像通道数翻倍，宽高减半，实现了类似于CNN中的池化操作。后续重复操作，数据依次送入Stage2、Stage3和Stage4，最终实现多尺度的特征提取。

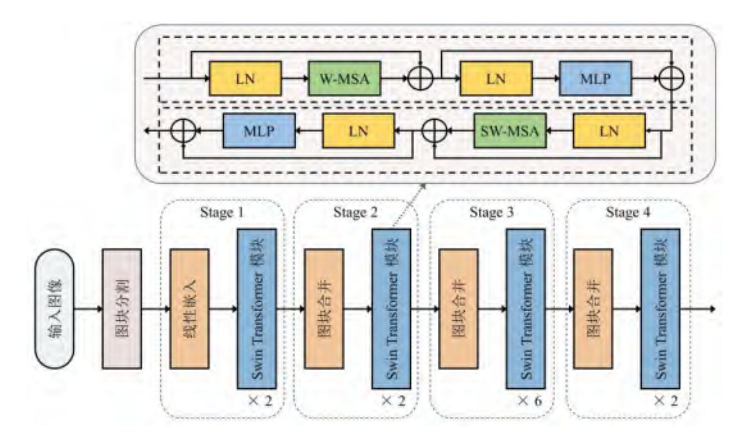


图4 Swin-Transformer模型架构

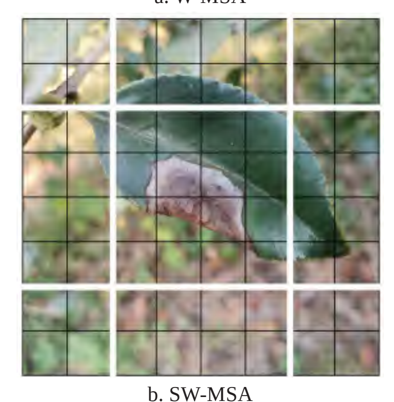
Fig. 4 Structure of Swin-Transformer

Swin Transformer模块主要由多头自注意力（Multi-head Self-Attention, MSA）和多层感知机（Multi Layer Perceptron, MLP）组成，在两个连续的Swin Transformer模块中首先使用了窗口多头自注意力（Windows Multi-head Self-Attention, W-

MSA)将图块划分为若干窗口，仅在每个窗口内进行自注意力计算，W-MSA结构如图5a所示。接着使用移位窗口多头自注意力（Shifted Window Multi-head Self Attention, SW-MSA）以实现不同窗口间的全局信息交互，结果如图5b所示。



a. W-MSA



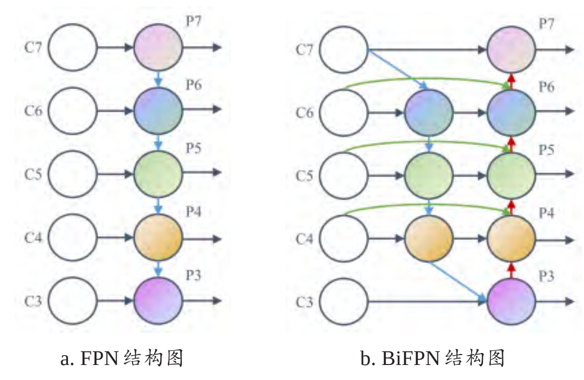
b. SW-MSA

图5 W-MSA和SW-MSA划分图像块示意图

Fig.5 Schematic diagram of W-MSA and SW-MSA divided image blocks

1.2.2 特征融合网络改进

深层特征图具有较强的语义信息，但分辨率低，细节感知能力差，适合大目标检测。浅层特征图感受野小，包含更多位置、信息细节，且几何表征能力强，适合小目标检测[20]。YOLACT使用的FPN网络[21]仅包含一条自顶向下的路径，用来将浅层特征传递到深层，如图6a所示。该算法侧重考虑深层特征对结果的影响，对浅层特征的利用不充分，甚至可能导致信息的丢失。BiFPN较FPN而言，采用了自顶向下和自底向上双向路径，将主干网络中不同尺度的特征直接与自底向上路径中的对应尺寸的特征融合，实现跨尺度连接，能够保留更多浅层特征信息，并在相同尺度的特征之间建立横向连接，以减轻由于网络层次过多而导致的特征信息丢失问题，提高检测精度。BiFPN结构如图6b所示。



a. FPN结构图 b. BiFPN结构图

图6 FPN与BiFPN结构图

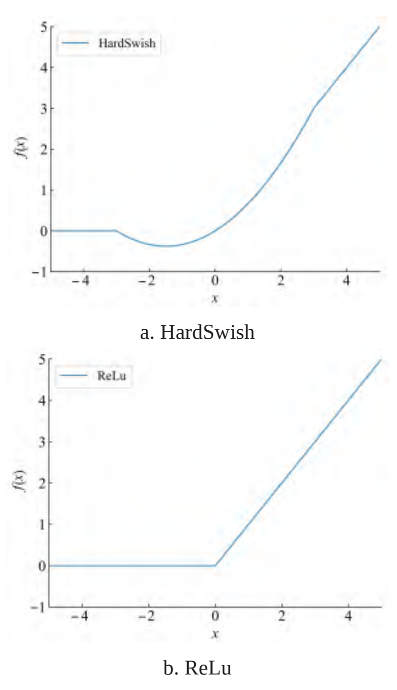
Fig. 6 Structure diagrams of FPN and BiFPN

1.2.3 激活函数优化

激活函数能够使神经网络学习和逼近更为复杂的非线性函数，加强了卷积神经网络对提取特征的表达[22]。本研究将原有的ReLu激活函数替换为在速度和精度等方面表现更优的HardSwish激活函数，表达如公式（2）所示。

$$\text{HardSwish}=\begin{cases}0 & x\leqslant -3\\ x & x\geqslant 3\\ \dfrac{x(x+3)}{6} & \text{otherwise}\end{cases}\tag{2}$$

函数图像如图7所示，在x>0的部分，ReLu是线性的，而HardSwish曲线更为平滑，在负值时仍存在较小梯度流，这使得训练更加容易收敛。



a. HardSwish

b. ReLu

图7 网络非线性研究ReLu和HardSwish函数对比

Fig.7 Comparison of ReLu and HardSwish functions in network nonlinearity research

2 实验与结果分析

2.1 实验环境

实验运行环境为Windows10操作系统，搭载主频2.9 GHz的Intel i5-9400F处理器，图形处理器为NVIDIA RTX 3070Ti 8 GB，内存32 GB，采用Pytorch 1.12.1深度学习框架，CUDA版本为11.3。为加快模型训练速度，特征提取主干网络部分均加载在ImageNet-1K上预训练的权重。训练过程采用随机梯度下降（Stochastic Gradient Descent, SGD）优化器，批次大小设置为8，训练轮次Epoch为300，动量参数设置为0.9，初始学习率为2e−3，学习率衰减策略选择余弦退火策略，权重衰减率为5e−4。

2.2 评价指标

本研究构建的油茶炭疽病数据集标注采用MS COCO标准格式，采用平均精确率均值（mean Average Precision, mAP）和平均召回率均值（mean Average Recall, mAR）作为实例分割掩模的评价指标，采用平均检测时间（td）评价模型性能。

mAP指所有类别AP的平均值，用来衡量多类别目标检测效果。在实例分割任务中，模型需要识别并分割出图像中的不同对象实例，因此需要对每个类别的分割结果进行精确评估。mAP通过计算每个类别的AP，并取其平均值，能够综合评估模型在各类别上的性能，从而更全面地反映模型的分割能力。mAP计算方法如公式（3）所示。mAR指所有类别AR的平均值，用来衡量多类别目标检测中模型找回所有相关实例的能力。在实例分割任务中，模型不仅需要准确分割出对象，还需要确保能够找到并分割出图像中的所有相关对象实例。因此，mAR作为评价指标，能够帮助评估模型在这方面的能力，确保模型能够尽可能地找到并分割出所有的对象实例。mAR计算方法如公式（4）所示。td直接反映了模型处理图像的速度，较短的检测时间意味着模型能够更高效地完成任务。

$$\mathrm{mAP} = \frac{\sum_{i=1}^{C} \mathrm{AP}_i}{C} \tag{3}$$

$$\mathrm{mAR} = \frac{\sum_{i=1}^{C} \mathrm{AR}_i}{C} \tag{4}$$

式中：C表示类别数；AP表示平均精确率，%；AR表示平均召回率，%。

本研究使用的mAP标准主要为mAP_{75}与mAP_{all}。mAP_{75}表示交并比（Intersection over Union, IoU）阈值为0.75时的平均精确率均值；mAP_{all}表示IoU阈值区间为0.5～0.95，步长为0.05下的平均精确率均值。其中IoU表示预测框和真实框的重合程度，用于衡量物体检测结果与真实值匹配的效果。IoU计算方法如公式（5）所示。

$$\mathrm{IoU} = \frac{A \cap B}{A \cup B} \tag{5}$$

式中：A表示预测框面积，pixel；B表示真实框面积，pixel。

2.3 消融实验

改进后的YOLACT模型使用Swin-Transformer作为特征提取主干网络，引入BiFPN实现双向特征融合，采用HardSwish激活函数增强网络非线性能力。为评估各项改进策略的有效性，在划分的测试集上保持相同实验条件进行消融实验，结果如表2所示。

表2　改进的YOLACT算法消融实验结果

Table 2 Results of ablation experiments with the improved YOLACT

编号	主干网络	BiFPN	HardSwish	mAP_{75}/%	mAP_{all}/%	mAR/%
0	ResNet50	×	×	81.1	75.8	83.7
1	Swin-T	×	×	86.7	78.0	91.2
2	Swin-T	√	×	86.5	78.1	91.3
3	Swin-T	√	√	86.8	78.3	91.6

注：×表示不使用该项改进因素；√表示使用该项改进因素。

由表2可知，模型0为原YOLACT模型，作为基线模型。模型1使用Swin-Transformer作为特征提取主干网络，mAP_{75}、mAP_{all}和mAR分别为86.7%、78.0%以及91.2%，较原模型相比分别提升5.6%、2.2%和7.5%。油茶炭疽病有着病斑大小、形态各异的特点，Swin-Transformer凭借自注意力机制对全局特征信息有更强的提取能力，能兼顾全局和局部的特征信息，表现更优。模型2在模型1基础上又引入了BiFPN，mAP_{75}、mAP_{all}和mAR为86.5%、78.1%和91.3%，与模型1相比分别下降了0.2%、提升0.1%以及提升0.1%。其中mAP_{75}略有下降，这可能由于mAP_{75}主要关注IoU阈值为0.75时的性能，这是一个相对较高的阈值，要求预测框与真实框之间有很高的重叠度。BiFPN更注重多尺度特征的融合，但在某些情况下可能没有FPN那么精确地在高IoU阈值下匹配目标。因此，mAP_{75}会略有下降。而mAP_{all}则考虑了多个IoU阈值下的平均精确率，包括一些较低的阈值，由于双向路径的BiFPN增进了不同层级间的特征融合，使得更多的

特征信息得以保留，因此整体性能会因BiFPN的特征融合能力而提升。模型3在模型2基础上使用了HardSwish激活函数，其mAP_{75}、mAP_{all}和mAR分别为86.8%、78.3%和91.6%，与模型2相比分别提升0.3%、0.2%和0.3%。HardSwish可以增强网络非线性性能力且具有更好的鲁棒性，提高了神经网络的表达能力，有效缓解ReLu神经元坏死的问题。Camellia-YOLACT模型与原模型相比mAP_{75}提升5.7%、mAP_{all}提升了2.5%，mAR提升了7.9%，结果表明改进方案具有有效性。

2.4 与其他分割方法对比

为进一步验证Camellia-YOLACT模型的检测性能，选取常用的实例分割模型Mask R-CNN以及SOLO[23]进行比较。Mask R-CNN与SOLO使用与上述相同数据集，训练至收敛。图8为不同算法分割测试效果，表3为对比实验结果。



a. 原图 b. Camelia-YOLACT c. Mask R-CNN d. SOLO

图8 不同算法对油茶炭疽病分割效果

Fig.8 Effectiveness of different algorithms for segmentation of *Camellia oleifera* leaf anthracnose

表3 不同算法的油茶炭疽病分割效果对比实验

Table 3 Comparative experiment on the segmentation effect of different algorithms for *Camellia oleifera* leaf anthracnose

模型	mAP_{75}/%	mAP_{all}/%	mAR/%	td/ms
YOLACT	81.1	75.8	83.7	46.85
Mask R-CNN	90.4	81.5	92.3	168.16
SOLO	87.1	77.8	90.4	75.73
Camellia-YOLACT	86.8	78.3	91.6	53.50

如表3所示，Camellia-YOLACT在精度方面表现出优于原模型的性能，虽然由于特征提取主干网络的优化导致参数量的提升，使得检测时间略有增加，但这并未影响模型在实时性方面的优势。对比同为单阶段实例分割模型的SOLO，mAP_{75}指标降低0.3%，mAP_{all}提高0.5%，mAR则领先1.2%，检测时间提升22.23 ms。此外，与精度高但推理速度较慢的Mask R-CNN相比，Camellia-YOLACT检测速度提升了2倍，确保了油茶叶部炭疽病分割速度的实时性要求。

具体来说，实时性高的优点使得Camellia-YOLACT能够在茶园中实现快速、准确的病斑分割。茶园中的环境复杂多变，油茶叶片的病害症状也可能多种多样，因此需要一个能够快速响应的模型。Camellia-YOLACT的高实时性能够满足这一需求，使得农民或管理人员能够实时获取信息，并采取相应的防治措施，从而有效控制病害的扩散，保障油茶树的健康生长。

2.5 油茶炭疽病分级实验

为验证所提出的Camellia-YOLACT模型在自然环境下对油茶炭疽病严重程度分级的效果，于2023年11月22在江西农业大学油茶种植基地进行室外实验，如图9所示，分别使用便携式计算机和Intel RealSense D455视觉相机作为上位机和图像采集设备，对油茶炭疽病严重程度进行分级实验。

选取36个不同程度的油茶炭疽病叶片，保留叶片样本图像进行人工像素标注并计算K值，对比模型预测结果与人工分级的结果，实验结果如表4所示，可知本研究所提出的Camellia-YOLACT模型对油茶炭疽病严重程度的分级结果良好，正确分级个数34个，分级正确率为94.4%，K值平均绝对误差为1.09%。分级有误的几组实验，通过分析图像，发现图像中均存在多个大小各异的病斑，其原因是分割的掩膜边缘不够精准，导致误差增大，在未来的研究中模型还需进一步优化。

3 结　论

为克服传统人工判别油茶炭疽病严重程度的局限性，通过引入智能化的分级技术，提高病害评估的准确性和效率，本研究提出了一种基于改进YOLACT的分割模型Camellia-YOLACT，该模型能够精准分割病斑进而实现对油茶炭疽病的严重程度自动分级。主要结论如下。

1）在YOLACT的特征提取主干网络部分引入



a. 实验设备

b. 实验效果

图9　油茶叶片炭疽病严重程度分级实验

Fig.9 *Camellia oleifera* leaf anthracnose injection severity grading experiment

Swin-Transformer。使用BiFPN替换了原始的FPN，双向特征融合通路使得不同尺度的特征信息能够更好地融合和利用，提升了模型的特征融合能力。HardSwish的使用增强了模型的非线性表达能力。经实验表明，Camellia-YOLACT模型 mAP_{75} 为86.8%，较原模型提升了5.7%；mAP_{all} 为78.3%，提升了2.5%；mAR为91.6%，提升了7.9%。对比其他分割模型，Camellia-YOLACT模型在精度和速度方面综合表现更优。

2）通过Camellia-YOLACT模型分割病斑区域和叶片区域，计算病斑区域面积除以叶片面积得到病害严重程度K。经实验表明，该方法分级正确率为94.4%，K值的平均绝对误差为1.09%。能够有效完成自然环境下油茶炭疽病的严重程度分级。

尽管本算法在综合性能体现出一定优势，但在某些特定场景下，其准确性和稳定性仍有待提升。例如，当叶片上小尺寸病斑较多时，病斑分割的精度会降低。后续可增加训练集中小目标病斑样本的数量，进一步提高分割精度。在室外实验中遇到叶片有部分重叠遮挡时，分割效果也会受到影响，未来研究需进一步考虑油茶炭疽病害叶片有遮挡时的分割。考虑检测的便捷性，后续可将Camellia-YOLACT油茶叶部炭疽病害检测模型植入嵌入式系

表4　Camellia-YOLACT算法油茶炭疽病分级实验

Table 4 Experimental classification of *Camellia oleifera* anthracnose by Camellia-YOLACT method

编号	真实值		预测值		K绝对误差/%	编号	真实值		预测值		K绝对误差/%
	K/%	等级	K/%	等级			K/%	等级	K/%	等级	
1	58.31	3	57.28	3	1.03	19	80.68	4	77.11	4	3.57
2	43.14	2	43.05	2	0.09	20	23.92	1	21.31	1	2.61
3	0.00	0	0.00	0	0.00	21	6.26	1	6.24	1	0.02
4	7.32	1	6.84	1	0.48	22	16.55	1	13.69	1	2.86
5	26.08	2	24.56	1	1.52	23	22.32	1	20.28	1	2.04
6	8.54	1	8.03	1	0.51	24	7.83	1	7.04	1	0.79
7	63.88	3	63.03	3	0.85	25	21.80	1	21.50	1	0.30
8	10.39	1	9.34	1	1.05	26	0.00	0	0.00	0	0.00
9	52.65	3	49.12	2	3.53	27	8.18	1	7.67	1	0.51
10	76.44	4	75.34	4	1.10	28	12.97	1	10.95	1	2.02
11	53.34	3	53.26	3	0.08	29	19.50	1	18.38	1	1.12
12	0.00	0	0.00	0	0.00	30	13.77	1	13.55	1	0.22
13	0.00	0	0.00	0	0.00	31	28.68	2	27.64	2	1.04
14	12.19	1	11.29	1	0.90	32	7.21	1	4.50	1	2.71
15	18.18	1	17.99	1	0.20	33	14.81	1	13.07	1	1.74
16	47.29	2	46.20	2	1.09	34	5.04	1	4.72	1	0.32
17	62.31	3	59.92	3	2.39	35	11.06	1	10.50	1	0.56
18	36.39	2	35.86	2	0.53	36	18.22	1	16.85	1	1.37

注：K值为0.00，代表没有检测出病斑；等级为0，代表健康。

统或可移动终端，以充分发挥其高实时性的优势，以期实现油茶叶片图像数据的实时分析，病害监测的自动化，并向用户发送病害预警信息，最终达到油茶叶部炭疽病害检测的智能化。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献：

[1] 张立伟, 王辽卫. 我国油茶产业的发展现状与展望[J]. 中国油脂, 2021, 46(6): 6-9, 27.
ZHANG L W, WANG L W. Prospect and development status of oil-tea camellia industry in China[J]. China oils and fats, 2021, 46(6): 6-9, 27.

[2] 吴鹏飞, 姚小华. 种植密度对普通油茶炭疽病病害发生的影响[J]. 中国油料作物学报, 2019, 41(3): 455-460.
WU P F, YAO X H. Effect of planting density on anthracnose occurrence of *Camellia oleifera*[J]. Chinese journal of oil crop sciences, 2019, 41(3): 455-460.

[3] 张蕊, 李锦涛. 基于深度学习的场景分割算法研究综述[J]. 计算机研究与发展, 2020, 57(4): 859-875.
ZHANG R, LI J T. A Survey on Algorithm Research of Scene Parsing Based on Deep Learning[J]. Journal of computer research and development, 2020, 57(4): 859-875.

[4] HAQUE M A, MARWAHA S, ARORA A, et al. A lightweight convolutional neural network for recognition of severity stages of maydis leaf blight disease of maize[J]. Frontiers in plant science, 2022, 13: ID 1077568.

[5] PRABHAKAR M, PURUSHOTHAMAN R, AWASTHI D P. Deep learning based assessment of disease severity for early blight in tomato crop[J]. Multimedia tools and applications, 2020, 79(39): 28773-28784.

[6] TENDANG S, CHAMNONGTHAI K. Rice-disease severity level estimation using deep convolutional neural network[C]// 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications (ITC-CSCC). Piscataway, New Jersey, USA: IEEE, 2021: 1-4.

[7] 万军杰, 祁力钧, 卢中奥, 等. 基于迁移学习的GoogLeNet果园病虫害识别与分级[J]. 中国农业大学学报, 2021, 26(11): 209-221.
WAN J J, QI L J, LU Z A, et al. Recognition and grading of diseases and pests in orchard by GoogLeNet based on transfer learning[J]. Journal of China agricultural university, 2021, 26(11): 209-221.

[8] LIU B, DING Z, TIAN L, et al. Grape leaf disease identification using improved deep convolutional neural networks[J]. Frontiers in plant science, 2020, 11: ID 1082.

[9] 王振, 张善文, 赵保平. 基于级联卷积神经网络的作物病害叶片分割[J]. 计算机工程与应用, 2020, 56(15): 242-250.
WANG Z, ZHANG S W, ZHAO B P. Crop diseases leaf segmentation method based on cascade convolutional neural network[J]. Computer engineering and applications, 2020, 56(15): 242-250.

[10] GARG K, BHUGRA S, LALL B. Automatic quantification of plant disease from field image data using deep learning[C]// 2021 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway, New Jersey, USA: IEEE, 2021: 1965-1972.

[11] GONÇALVES JULIANO P, PINTO FRANCISCO A C, QUEIROZ DANIEL M, et al. Deep learning architectures for semantic segmentation and automatic estimation of severity of foliar symptoms caused by diseases or pests[J]. Biosystems engineering, 2021, 210: 129-142.

[12] 茹佳棋, 吴斌, 翁翔, 等. 基于改进UNet++模型的葡萄黑腐病病斑分割和病害程度分级[J]. 浙江农业学报, 2023, 35(11): 2720-2730.
RU J Q, WU B, WENG X, et al. Disease spot segmentation and disease degree classification of grape black rot based on improved UNet++ model[J]. Acta agriculturae Zhejiangensis, 2023, 35(11): 2720-2730.

[13] 邓朝, 纪苗苗, 任永泰. 基于Mask R-CNN的马铃薯叶片晚疫病量化评价[J]. 扬州大学学报(农业与生命科学版), 2022, 43(1): 135-142.
DENG Z, JI M M, REN Y T. Quantitative evaluation of potato late blight disease based on Mask R-CNN[J]. Journal of Yangzhou university (agricultural and life science edition), 2022, 43(1): 135-142.

[14] LIU Z, LIN Y T, CAO Y, et al. Swin Transformer: Hierarchical Vision Transformer using Shifted Windows[C]// 2021 IEEE/CVF International Conference on Computer Vision (ICCV). Piscataway, New Jersey, USA: IEEE, 2021: 10012-10022.

[15] TAN M X, PANG R M, LE Q V. EfficientDet: scalable and efficient object detection[C]// 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, New Jersey, USA: IEEE, 2020: 10781-10790.

[16] 安徽省市场监督管理局. 茶炭疽病测报调查与防治技术规程: DB34/T 3863—2021[S].

[17] SHU J H, NIAN F D, YU M H, et al. An improved mask R-CNN model for multiorgan segmentation[J]. Mathematical Problems in Engineering, 2020, 2020: 1-11.

[18] ZHANG Z, HUANG S, LIU X, et al. Adversarial attacks on YOLACT instance segmentation[J]. Computers & Security, 2022, 116: ID 102682.

[19] SHAFIQ M, GU Z. Deep residual learning for image recognition: A survey[J]. Applied sciences, 2022, 12(18): ID 8972.

[20] 杨毅, 桑庆兵. 多尺度特征自适应融合的轻量化织物瑕疵检测[J]. 计算机工程, 2022, 48(12): 288-295.
YANG Y, SANG Q B. Lightweight-fabric defect detection based on adaptive fusion of multiscale features[J]. Computer engineering, 2022, 48(12): 288-295.

[21] ZHU L, LEE F, CAI J, et al. An improved feature pyramid network for object detection[J]. Neurocomputing, 2022, 483: 127-139.

[22] 蓝金辉, 王迪, 申小盼. 卷积神经网络在视觉图像检测的研究进展[J]. 仪器仪表学报, 2020, 41(4): 167-182.
LAN J H, WANG D, SHEN X P. Research progress on visual image detection based on convolutional neural network[J]. Chinese journal of scientific instrument, 2020, 41 (4): 167-182.

[23] WANG X, KONG T, SHEN C, et al. Solo: Segmenting objects by locations[C]// Computer Vision-ECCV 2020. ECCV 2020. Lecture Notes in Computer Science. Cham, Germany: Springer, 2020: 649-665.

Severity Grading Model for *Camellia Oleifera* Anthracnose Infection Based on Improved YOLACT

NIE Ganggang[1,2], RAO Honghui[1,2*], LI Zefeng[1,2], LIU Muhua[1,2]

（1. *College of Engineering, Jiangxi Agriculture University, Nanchang* 330045, *China*;
2. *Jiangxi Key Laboratory of Modern Agricultural Equipment, Jiangxi Agriculture University, Nanchang* 330045, *China*）

Abstract:
[Objective] *Camellia oleifera* is one of the four major woody oil plants in the world. Diseases is a significant factor leading to the decline in quality of *Camellia oleifera* and the financial loss of farmers. Among these diseases, anthracnose is a common and severe disease in *Camellia oleifera* forests, directly impacting yields and production rates. Accurate disease assessment can improve the prevention and control efficiency and safeguarding the farmers' profit. In this study, an improved You Only Look at CoefficienTs (YOLACT) based method was proposed to realize automatic and efficient grading of the severity of *Camellia oleifera* leaf anthracnose.
[Methods] High-resolution images of *Camellia oleifera* anthracnose leaves were collected using a smartphone at the National *Camellia oleifera* Seed Base of Jiangxi Academy of Forestry, and finally 975 valid images were retained after a rigorous screening process. Five data enhancement means were applied, and a data set of 5 850 images was constructed finally, which was divided into training, validation, and test sets in a ratio of 7:2:1. For model selection, the Camellia-YOLACT model was proposed based on the YOLACT instance segmentation model, and by introducing improvements such as Swin-Transformer, weighted bi-directional feature pyramid network, and HardSwish activation function. The Swin Transformer was utilized for feature extraction in the backbone network part of YOLACT, leveraging the global receptive field and shift window properties of the self-attention mechanism in the Transformer architecture to enhance feature extraction capabilities. Additionally, a weighted bidirectional feature pyramid network was introduced to fuse feature information from different scales to improve the detection ability of the model for objects at different scales, thereby improving the detection accuracy. Furthermore, to increase the the model's robustness against the noise in the input data, the HardSwish activation function with stronger nonlinear capability was adopted to replace the ReLu activation function of the original model. Since images in natural environments usually have complex background and foreground information, the robustness of HardSwish helped the model better handling these situations and further improving the detection accuracy. With the above improvements, the Camellia-YOLACT model was constructed and experimentally validated by testing the *Camellia oleifera* anthracnose leaf image dataset.
[Results and Discussions] A transfer learning approach was used for experimental validation on the *Camellia oleifera* anthracnose severity grading dataset, and the results of the ablation experiments showed that the mAP_{75} of Camellia-YOLACT proposed in this study was 86.8%, mAP_{all} was 78.3%, mAR was 91.6% which were 5.7%, 2.5% and 7.9% higher than YOLACT model. In the comparison experiments, Camellia-YOLACT performed better than Segmenting Objects by Locations (SOLO) in terms of both accuracy and speed, and its detection speed was doubled compared to Mask R-CNN algorithm. Therefore, the Camellia-YOLACT algorithm was suitable in *Camellia oleifera* gardens for anthracnose real-time segmentation. In order to verify the outdoors detection performance of Camellia-YOLACT model, 36 groups of *Camellia oleifera* anthracnose grading experiments were conducted. Experimental results showed that the grading correctness of *Camellia oleifera* anthracnose injection severity reached 94.4%, and the average absolute error of K-value was 1.09%. Therefore, the Camellia-YOLACT model proposed in this study has a better performance on the grading of the severity of *Camellia oleifera* anthracnose.
[Conclusions] The Camellia-YOLACT model proposed got high accuracy in leaf and anthracnose segmentation of *Camellia oleifera*, on the basis of which it can realize automatic grading of the severity of *Camellia oleifera* anthracnose. This research could provide technical support for the precise control of *Camellia oleifera* diseases.

Key words: *Camellia oleifera*; leaf disease; anthracnose; BiFPN; YOLACT; Transformer; deep learning

Foundation items: Jiangxi Science and Technology Planning Project(20141BBF60057); Provincial Forestry Bureau Camellia Fruit Research Special Project (YCYJZX2023221)

Biography: NIE Ganggang, E-mail: nieganggang1997@163.com

***Corresponding author:** RAO Honghui, E-mail: rhh58@sohu.com